Hypervirulent *emm59* Clone in Invasive Group A *Streptococcus* Outbreak, Southwestern United States

David M. Engelthaler,[1] Michael Valentine,[1] Jolene Bowers, Jennifer Pistole, Elizabeth M. Driebe, Joel Terriquez, Linus Nienstadt, Mark Carroll, Mare Schumacher, Mary Ellen Ormsby, Shane Brady, Eugene Livar, Del Yazzie, Victor Waddell, Marie Peoples, Kenneth Komatsu, Paul Keim

The hyper-virulent *emm*59 genotype of invasive group A *Streptococcus* was identified in northern Arizona in 2015. Eighteen isolates belonging to a genomic cluster grouped most closely with recently identified isolates in New Mexico. The continued transmission of *emm*59 in the southwestern United States poses a public health concern.

Several cases of invasive group A *Streptococcus* (GAS) disease were detected in January 2015 in a northern Arizona hospital. A substantive percentage of the cases were associated with a homeless shelter and a local jail; outbreak case-patients were predominantly male and Native American. Other studies have shown an increase in infection risk for invasive GAS in Native American/First Nations populations (*1*,*2*), and outbreaks within this population in Arizona have been previously documented (*3*). Whole genome sequence analysis determined that the hypervirulent subtype *emm*59 was present among the first cases analyzed in early 2015. *emm*59 is known to have caused a nationwide outbreak of invasive GAS in Canada during 2006–2009 (*4,5*), and cases and outbreaks have been reported in the United States (*6*–*8*).

The Study

We identified isolates for sequencing from 29 invasive GAS cases diagnosed in patients in a northern Arizona hospital during January–July 2015 and randomly selected an additional 99 GAS isolates from a repository of >2,000 Arizona GAS isolates collected during 2002–2006 (no isolates from patients in Arizona were available for 2007–2014). Four additional isolates from central Arizona identified in 2015 were included in the analysis (online Technical Appendix Table, http://wwwnc.cdc.gov/EID/article/22/1/15-1200-Techapp1.pdf). All isolates were grown on 5% sheep blood tryptic soy agar plates (Hardy Diagnostics, Santa Maria, CA), and incubated at 37°C with 5% CO_2. DNA was extracted by using a DNeasy Blood and Tissue Kit (QIAGEN, Valencia, CA, USA) following manufacturer's protocol. Genomic DNA libraries were prepared by using the Nextera XT library prep kit (Illumina, San Diego, CA) and sequenced with paired-end reads (250 bp) on an Illumina MiSeq instrument, as previously described (*9*). The finished genome of the *emm*59 Canadian clone MGAS15252 (GenBank accession no. CP003116) and high-quality publicly available sequence-read data from 44 US isolates, from NCBI short read archive (BioProject #PRJNA194066), were included in the subsequent phylogenetic analyses. The final core genome (all nucleotide loci found in all genomes) for single-nucleotide polymorphism (SNP) detection was 1,636,024 bp (98.6% of reference).

We used NASP SNP analysis pipeline (http://tgen-north.github.io/NASP/) for whole-genome SNP typing as previously described (*10*). SNP matrices were developed for both the whole species and the *emm*59-only analyses. We used MEGA version 5.2.2 software (*11*) to generate maximum parsimony phylogenetic trees. Regions of high SNP density were identified as possible regions of recombination and were further analyzed for impact on the consistency index. Genomes were assembled by using UGAP (https://github.com/jasonsahl/UGAP). GAS *emm* subtypes were assigned by using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi), querying the study genome assemblies against the Centers for Disease Control and Prevention's (CDC) *emm* type-specific sequence database (http://www.cdc.gov/streplab/m-proteingene-typing.html). We resolved dual *emm*-type hits using CDC's *emm* typing Sanger sequencing primers (http://www.cdc.gov/streplab/protocol-emm-type.html) as a BLAST query and noting hit locations.

We identified 18 of the 29 contemporary northern Arizona isolates as subtype *emm*59; the remaining isolates

Author affiliations: Translational Genomics Research Institute, Flagstaff, Arizona, USA (D. Engelthaler, M. Valentine, J. Bowers, E.M. Driebe, P. Keim); Arizona Department of Health Services, Phoenix, Arizona, USA (J. Pistole, S. Brady, E. Livar, V. Waddell, K. Komatsu); Northern Arizona Healthcare, Flagstaff (J. Terriquez, L. Nienstadt, M. Carroll); Coconino County Public Health Services District, Flagstaff (M. Schumacher, M.E. Ormsby, M. Peoples); Navajo Division of Health, Window Rock, Arizona, USA (D. Yazzie); Northern Arizona University, Flagstaff (P. Keim)

DOI: http://dx.doi.org/10.3201/eid2204.151582

[1]These authors contributed equally to this article.

were composed of 6 additional *emm* types: *emm*1 (n = 2), *emm*5 (n = 2), *emm*58 (n = 1), *emm*81 (n = 2), *emm*83 (n = 1), *emm*89 (n = 2), and *emm*94 (n = 1). The 99 historical and 4 contemporary background Arizona isolates included 25 distinct *emm* types (online Technical Appendix Table). No *emm*59 isolates were identified in this background set, and none had been previously reported in Arizona.

The 18 Arizona *emm*59 cases occurred during January–July 2015 (Table). An *emm*59-only phylogenetic analysis demonstrated the apparent presence of multiple lineages of *emm*59 in the 2015 Arizona isolates (Figure 1). A distinct clone consisting of 14 of the 18 *emm*59 isolates were separated from each other by only 0–4 SNPs, genomically supporting the presence of an ongoing outbreak; ≥8 of these patients were epidemiologically linked to physical contact, cohabitation, or both with 1 other person (data not shown). The additional *emm*59 isolates make up additional lineages separated from one other by 8–28 SNPs. No recombination was identified among the Arizona isolates. A relatively large number of SNPs and indels were seen within an approximate 23-kilobase region (Figure 1). This region has been previously reported to contain mutational hotspots associated with virulence (*12*,*13*). Considering the presumptive positive selective force on this region, SNPs within the region were not included in the final phylogenetic analysis.

Table. Epidemiologic data for 18 case-patients with invasive *emm*59 group A *Streptococcus* infection, Arizona, USA, 2015*

Category	Value
Race	
American Indian or Alaskan Native	15 (83)
White	3 (7)
Sex	
F	4 (22)
M	14 (88)
Mean age, y (range)	40 (26–79)
Clinical information	
Cellulitis	7 (39)
Necrotizing fasciitis	5 (28)
Sepsis	9 (50)
Risk factors	
Injury	7 (39)
Alcohol abuse	10 (56)
Homeless	8 (44)
Living in shelter	5 (28)
Local jail term within ≈1 mo. of diagnosis	6 (33)

*Values are no. (%) patients except as indicated. Epidemiologic data based on available information.

When compared with all other publicly available US *emm*59 isolate genomes, nearly all the genomes identified in the United States were closely related to each other and to the Canadian clone MGAS15252; individual isolate SNP branch lengths ranged from 0 to 10 (Figure 2). The Arizona outbreak isolates were separated from 2 New Mexico isolates by 4 and 5 SNPs each; these isolates fell within the overall Arizona clade and were subsequently included

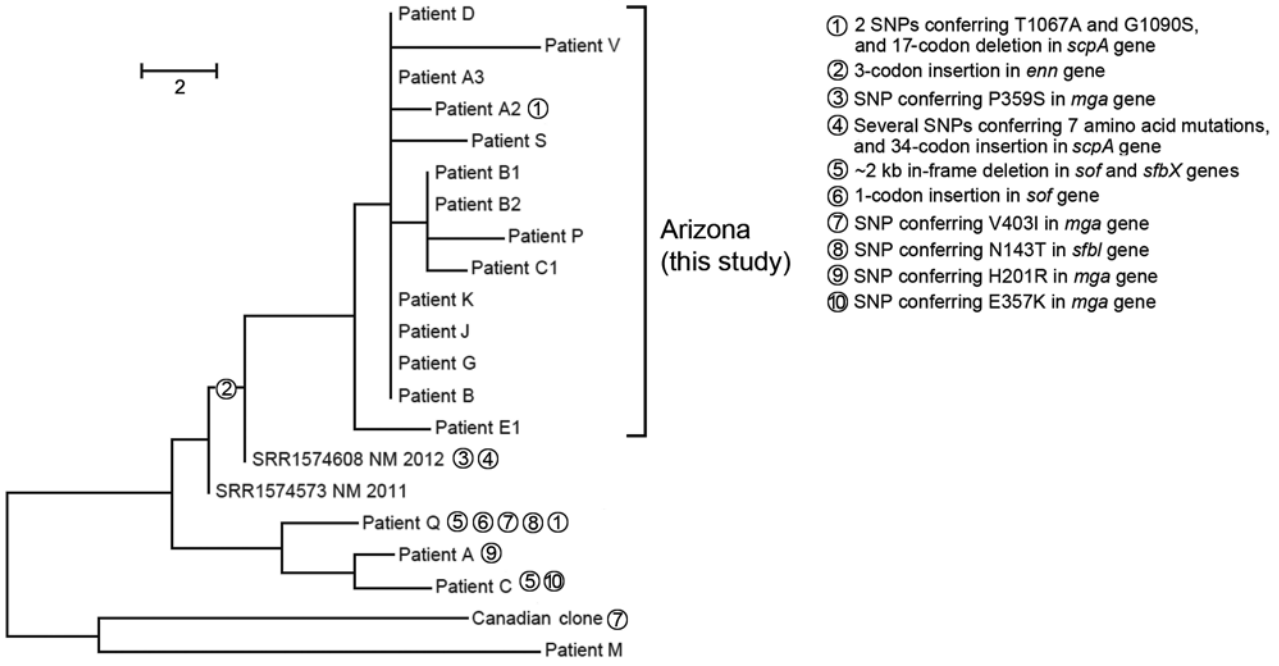

Figure 1. Phylogenetic single-nucleotide polymorphism (SNP) tree of *emm*59 isolates from a northern Arizona hospital displaying distribution of mutations in a 23kb positively selected region during invasive group A *Streptococcus* outbreak, southwestern United States. Maximum parsimony tree of all SNP loci (n = 58) in *emm*59 isolates (n = 18) from Arizona, 2 recent New Mexico isolate genomes, and the Canadian clone reference isolate MGAS15252. Consistency index = 1.0. Branch lengths represent numbers of SNPs between isolates; unit bar is in the figure. Numbered circles distinguish lineages of selected mutations in *scpA, enn, sfbI, mga, sfbx,* and *sof* genes in a 23-kb hotspot mutational region. Scale bar indicates SNPs.

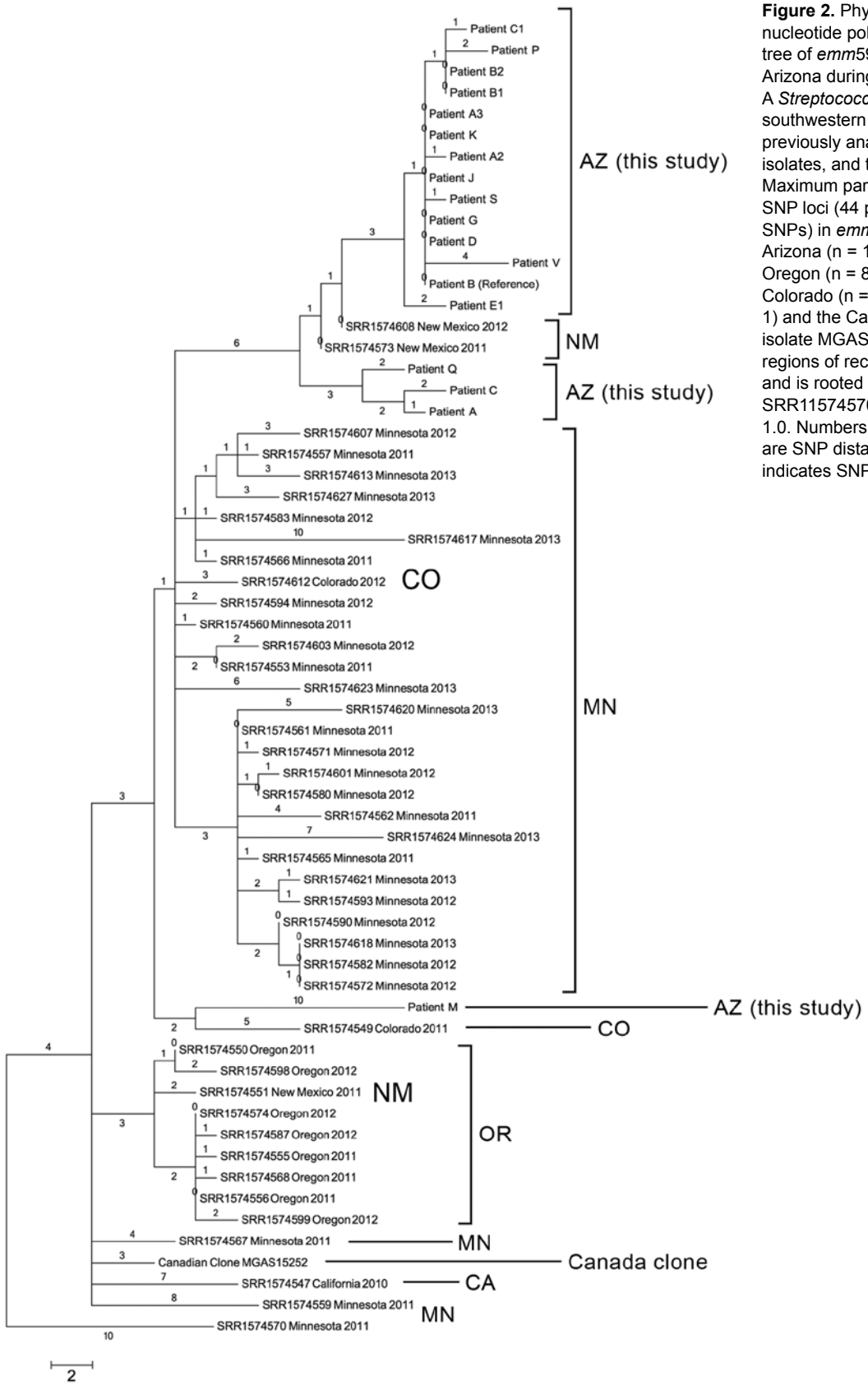

Figure 2. Phylogenetic single-nucleotide polymorphism (SNP) tree of *emm*59 isolates from Arizona during invasive group A *Streptococcus* outbreak in the southwestern United States, previously analyzed US *emm*59 isolates, and the Canadian clone. Maximum parsimony tree of all 177 SNP loci (44 parsimony informative SNPs) in *emm*59 isolates from Arizona (n = 18), Minnesota (n = 29), Oregon (n = 8), New Mexico (N = 3), Colorado (n = 2), and California (n = 1) and the Canadian clone reference isolate MGAS15252. Tree has regions of recombination removed and is rooted with Minnesota isolate SRR11574570. Consistency index = 1.0. Numbers above branches are SNP distances. Scale bar indicates SNPs.

in the Arizona-only phylogenetic analysis (Figure 1). Conversley, the isolate from patient M appears more distant from the larger Arizona population. The Arizona clades, with the exception of that of the isolate from patient M, all appear to arise from the large Minnesota polytomy. The previously estimated 1.3–2.1 SNPs/year mutation rates for GAS (*14*,*15*) further support the Arizona outbreak as being caused by a single clone, likely originating from New Mexico and being spread over 6–12 months.

Conclusions

The *emm*59 subtype of GAS, the etiologic agent of a substantial nationwide outbreak of invasive GAS in Canada during 2006–2009 (*4*), is now present in Arizona, causing at least 1 outbreak of epidemiologically and genomically linked cases and several additional epidemiologically unrelated cases. The lack of *emm*59 in background isolates in Arizona from the previous decade, along with its low genetic diversity, suggests that *emm*59 emerged recently in Arizona. Following the *emm*59 epidemic in Canada, this subtype was subsequently seen in a few US states; a retrospective analyses of the Centers for Disease Control and Prevention Active Bacterial Core surveillance (ABCs) system (http://www.cdc.gov/abcs/reports-findings/survreports.pdf) identified 40 US *emm*59 isolates during 2000–2009 (*6*) and an additional 67 isolates during 2010–2012 (*7*). Of note, only 5 (of the 40 *emm*59 isolates from 2000–2009 (2 from Minnesota, 2 from California, and 1 from Oregon) appeared to be closely related to the Canadian clone (defined by the authors as being separated by <16 SNPs) (*6*); in contrast, all of the strains from the 2010–2012 survey appeared to be more closely related to the Canadian clone. The more recent ABCs analysis identified an increasing number of southwestern isolates, including 4 from Colorado and 6 from New Mexico (*7*), although no outbreaks were specifically described in these states (Arizona is not included in the ABCs system). Similar to this outbreak study, Olsen et al. (*7*), in an analysis of 60 MN *emm59* isolates from case-patients with identified race, determined that 25 (42%) were from Native Americans; of 5 isolates from New Mexico in that study, 3 were from Native Americans.

Given the apparent distal nature of the Arizona/New Mexico isolates to the Minnesota population in our study, it is reasonable to propose an unidentified epidemiologic relationship between these case populations. However, caution must be used in drawing conclusions regarding the relationships of isolates from disparate geographic regions because only limited comparable sequence data from previous *emm*59 studies in the United States (*7*) were publicly available to compare to the Arizona isolates. Epidemiologic investigations, along with healthcare provider and patient education activities, are ongoing in Arizona to further determine the extent of the current outbreak and the associated risk factors and to help mitigate effects and limit or prevent further spread to at-risk populations.

Dr. Engelthaler is an associate professor with the Translational Genomics Research Institute in Flagstaff, AZ. His research interests are in advancing epidemiology and clinical response through applied infectious disease genomics.

References

1. Hoge CW, Schwartz B, Talkington DF, Breiman RF, MacNeill EM, Englender SJ. The changing epidemiology of invasive group A streptococcal infections and the emergence of streptococcal toxic shock–like syndrome. A retrospective population-based study. JAMA. 1993;269:384–9. http://dx.doi.org/10.1001/jama.1993.03500030082037
2. Athey TB, Teatero S, Sieswerda LE, Gubbay JB, Marchand-Austin A, Li A, et al.. High incidence of invasive group A *Streptococcus* disease caused by strains of uncommon *emm* types in Thunder Bay, Ontario, Canada. J Clin Microbiol. 2016;54:83–92. http://dx.doi.org/10.1128/JCM.02201-15
3. Harris AM, Yazzie D, Antone-Nez R, Dinè-Chacon G, Kinlacheeny JB, Foley D, et al. Community-acquired invasive GAS disease among Native Americans, Arizona, USA, Winter 2013. Emerg Infect Dis. 2015;21:177–9. http://dx.doi.org/10.3201/eid2101.141148
4. Fittipaldi N, Beres SB, Olsen RJ, Kapur V, Shea PR, Watkins ME, et al. Full-genome dissection of an epidemic of severe invasive disease caused by a hypervirulent, recently emerged clone of group A *Streptococcus*. Am J Pathol. 2012;180:1522–34. http://dx.doi.org/10.1016/j.ajpath.2011.12.037
5. Tyrrell GJ, Lovgren M, Ibrahim Q, Garg S, Chui L, Boone TJ, et al. Epidemic of invasive pneumococcal disease, western Canada, 2005–2009. Emerg Infect Dis. 2012;18:733–40. http://dx.doi.org/10.3201/eid1805.110235
6. Fittipaldi N, Olsen RJ, Beres SB, Van Beneden C, Musser JM. Genomic analysis of *emm59* group A *Streptococcus* invasive strains, United States. Emerg Infect Dis. 2012;18:650–2. http://dx.doi.org/10.3201/eid1804.111803
7. Olsen RJ, Fittipaldi N, Kachroo P, Sanson MA, Long SW, Como-Sabetti KJ, et al. Clinical laboratory response to a mock outbreak of invasive bacterial infections: a preparedness study. J Clin Microbiol. 2014;52:4210–6. http://dx.doi.org/10.1128/JCM.02164-14
8. Brown CC, Olsen RJ, Fittipaldi N, Mormon ML, Fort PL, Neuwirth R, et al. Spread of virulent group A *Streptococcus* type *emm59* from Montana to Wyoming, USA. Emerg Infect Dis. 2014;20:679–81. http://dx.doi.org/10.3201/eid2004.130564
9. Driebe EM, Sahl JW, Roe CC, Bowers JR, Schupp JM, Gillece JD, et al. Using whole genome analysis to examine Recombination Across Diverse Sequence Types of *Staphylococcus aureus*. PLoS ONE. 2015;10:e0130955. http://dx.doi.org/10.1371/journal.pone.0130955
10. Struve C, Roe CC, Stegger M, Stahlhut SG, Hansen DS, Engelthaler DM, et al. Mapping the evolution of hypervirulent *Klebsiella pneumoniae.* MBio. 2015;6:e00630.
11. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using *Streptococcus*, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28:2731–9. http://dx.doi.org/10.1093/molbev/msr121
12. Fittipaldi N, Tyrrell GJ, Low DE, Martin I, Lin D, Hari KL, et al. Integrated whole-genome sequencing and temporospatial analysis of a continuing group A *Streptococcus* epidemic. Emerg Microbes Infect. 2013;2:e13. http://dx.doi.org/10.1038/emi.2013.13

13. Sanson M, O'Neill BE, Kachroo P, Anderson JR, Flores AR, Valson C, et al. A naturally occurring single amino acid replacement in multiple gene regulator of group A *Streptococcus* significantly increases virulence. Am J Pathol. 2015; 185:462–71.
14. Nasser W, Beres SB, Olsen RJ, Dean MA, Rice KA, Long SW, et al. Evolutionary pathway to increased virulence and epidemic group A *Streptococcus* disease derived from 3,615 genome sequences. Proc Natl Acad Sci U S A. 2014;111:E1768–76. http://dx.doi.org/10.1073/pnas.1403138111
15. Turner CE, Abbott J, Lamagnic T, Holden MTG, David S, Jones MD, Game L, Efstratiou A, Sriskandan S. Emergence of a new highly successful acapsular group A *Streptococcus* clade of the genotype emm89. mBio. 2015;6:e00622–15. http://dx.doi.org/10.1128/mBio.00622-15

Address for correspondence: David M. Engelthaler, TGen North, The Translational Genomics Research Institute, 3051 W. Shamrell Blvd., Suite 106, Flagstaff, AZ 86001, USA; email: dengelthaler@tgen.org

Article DOI: http://dx.doi.org/10.3201/eid2204.151582

Hypervirulent *emm*59 Clone in Invasive Group A *Streptococcus* Outbreak, Southwestern United States

Technical Appendix

Technical Appendix Table. Arizona Group A Streptococcus strain list.

Sample name	Isolation year	EmmType
Patient V	2015	emm59
Patient T	2015	emm58
Patient S	2015	emm59
Patient R	2015	emm89
Patient Q	2015	emm59
Patient P	2015	emm59
Patient O	2015	emm94
Patient N	2015	emm81
Patient M	2015	emm59
Patient L	2015	emm81
Patient K	2015	emm59
Patient J	2015	emm59
Patient I	2015	emm89
Patient H	2015	emm1
Patient G	2015	emm59
Patient F	2015	emm5
Patient E1	2015	emm59
Patient E	2015	emm5
Patient D	2015	emm59
Patient C1	2015	emm59
Patient C	2015	emm59
Patient B2	2015	emm59
Patient B1	2015	emm59
Patient B	2015	emm59
Patient A3	2015	emm59
Patient A2	2015	emm59
Patient A1	2015	emm83
Patient A	2015	emm59
Patient 50	2015	emm1
15-AZDH-31896	2015	emm1
15-AZDH-31642	2015	emm89
15-AZDH-31316	2015	emm89
15-AZDH-31315	2015	emm89
06-AZDH-917	2006	emm1
06-AZDH-889	2006	emm1
06-AZDH-6998	2006	emm83
06-AZDH-6944	2006	emm1
06-AZDH-6629	2006	emm1
06-AZDH-653	2006	emm1
06-AZDH-6182	2006	emm156
06-AZDH-6118	2006	emm80
06-AZDH-5439	2006	emm81
06-AZDH-5161	2006	emm1
06-AZDH-4827	2006	emm1
06-AZDH-4311	2006	emm1
06-AZDH-3955	2006	emm28
06-AZDH-390	2006	emm22
06-AZDH-3232	2006	emm75
06-AZDH-3204	2006	emm1
06-AZDH-3102	2006	emm1
06-AZDH-310013	2006	emm75

Sample name	Isolation year	EmmType
06-AZDH-299002	2006	emm1
06-AZDH-2950	2006	emm3
06-AZDH-2909	2006	emm118
06-AZDH-286011	2006	emm28
06-AZDH-284001	2006	emm28
06-AZDH-283016	2006	emm76
06-AZDH-281	2006	emm1
06-AZDH-2550	2006	emm12
06-AZDH-2484	2006	emm41
06-AZDH-1564	2006	emm118
06-AZDH-1120	2006	emm83
05-AZDH-9295	2005	emm33
05-AZDH-8840	2005	emm1
05-AZDH-8819	2005	emm12
05-AZDH-7607	2005	emm118
05-AZDH-7138	2005	emm22
05-AZDH-649	2005	emm83
05-AZDH-5954	2005	emm78
05-AZDH-5712	2005	emm22
05-AZDH-5225	2005	emm12
05-AZDH-4797	2005	emm1
05-AZDH-4700	2005	emm87
05-AZDH-3549	2005	emm94
05-AZDH-341	2005	emm6
05-AZDH-3409	2005	emm76
05-AZDH-3375	2005	emm12
05-AZDH-11701	2005	emm80
05-AZDH-1139	2005	emm12
05-AZDH-11329	2005	emm92
05-AZDH-11117	2005	emm81
05-AZDH-10477	2005	emm12
04-AZDH-8764	2004	emm83
04-AZDH-8624	2004	emm1
04-AZDH-8516	2004	emm6
04-AZDH-7857	2004	emm1
04-AZDH-7507	2004	emm89
04-AZDH-7482	2004	emm28
04-AZDH-7102	2004	emm75
04-AZDH-6995	2004	emm12
04-AZDH-6727	2004	emm156
04-AZDH-6516	2004	emm1
04-AZDH-6188	2004	emm28
04-AZDH-6011	2004	emm114
04-AZDH-5802	2004	emm58
04-AZDH-5600	2004	emm76
04-AZDH-5454	2004	emm80
04-AZDH-4661	2004	emm6
04-AZDH-4596	2004	emm28
04-AZDH-4517	2004	emm2
04-AZDH-3925	2004	emm1
03-AZDH-924	2003	emm2
03-AZDH-739	2003	emm83
03-AZDH-693	2003	emm12
03-AZDH-546	2003	emm3
03-AZDH-376	2003	emm5
03-AZDH-3258	2003	emm87
03-AZDH-3091	2003	emm1
03-AZDH-296	2003	emm1
03-AZDH-2839	2003	emm1
03-AZDH-2776	2003	emm1
03-AZDH-2617	2003	emm75
03-AZDH-2592	2003	emm3
03-AZDH-2592	2003	emm3
03-AZDH-2525	2003	emm89
03-AZDH-2235	2003	emm3
03-AZDH-1872	2003	emm114
03-AZDH-1822	2003	emm3
03-AZDH-1666	2003	emm2
03-AZDH-1622	2003	emm89

Sample name	Isolation year	EmmType
03-AZDH-2454	2003	emm1
03-AZDH-1105	2003	emm48
02-AZDH-273	2002	emm22
02-AZDH-268	2002	emm1
02-AZDH-249	2002	emm12
02-AZDH-238	2002	emm75
02-AZDH-221	2002	emm156
02-AZDH-198	2002	emm12
02-AZDH-181	2002	emm156
02-AZDH-172	2002	emm75
02-AZDH-166	2002	emm41
02-AZDH-158	2002	emm1
02-AZDH-154	2002	emm1